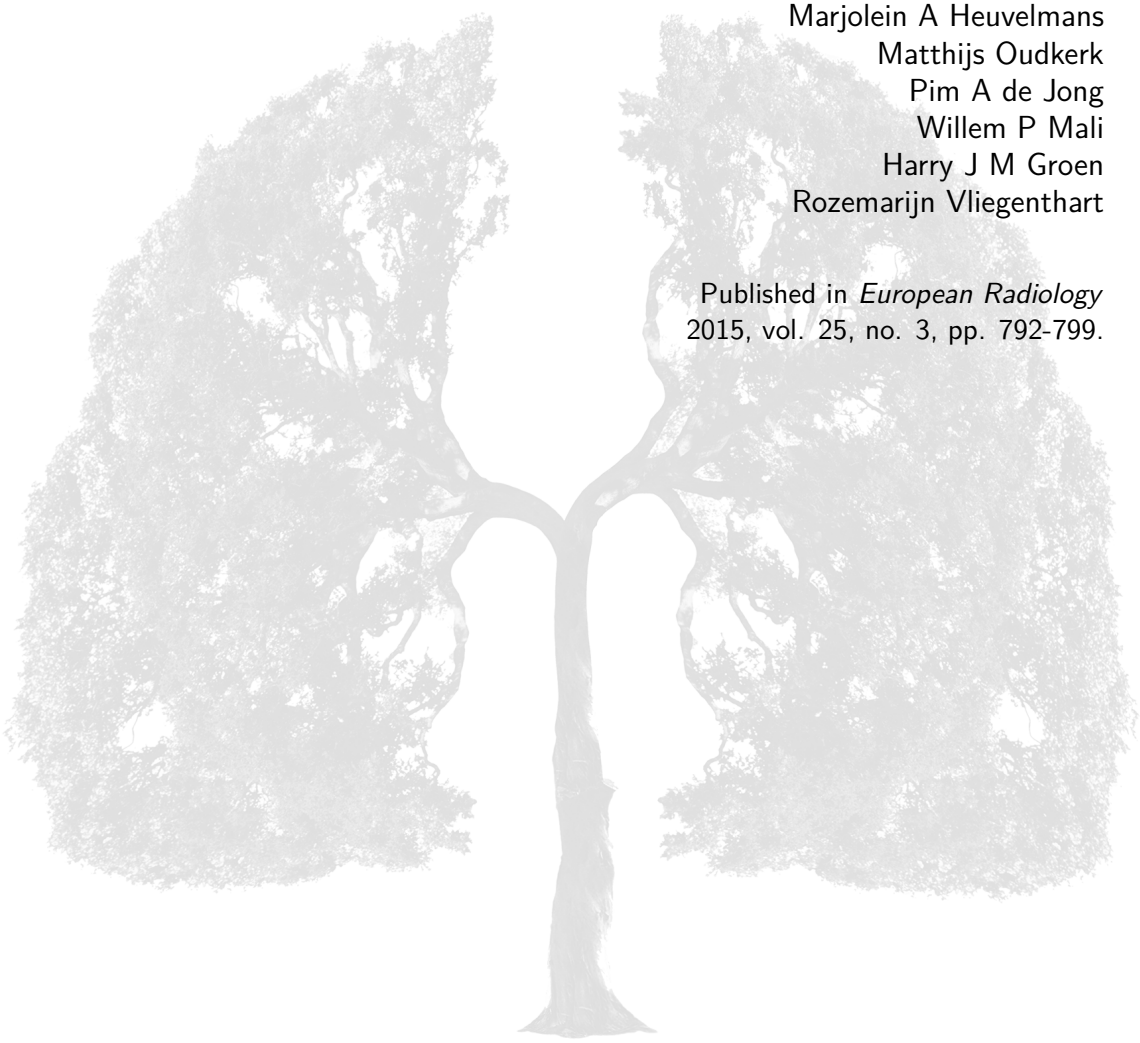
9

The impact of radiologists' expertise on screen results decisions in a CT lung cancer screening trial

Marjolein A Heuvelmans
Matthijs Oudkerk
Pim A de Jong
Willem P Mali
Harry J M Groen
Rozemarijn Vliegenthart

Published in *European Radiology*
2015, vol. 25, no. 3, pp. 792-799.



Abstract

Objective: To evaluate the impact of radiologists' expertise on screen result decisions made in a CT lung cancer screening trial.

Methods: In the NELSON-lung cancer screening trial, the baseline CT result was based on the largest lung nodule's volume. The protocol allowed radiologists to manually adjust screen result in case of high suspicion of benign or malignant nodule nature. In this study, all participants whose baseline CT result was based on a solid or part-solid nodule were included. Adjustments by radiologists at baseline were evaluated. Histology was the reference for diagnosis, or, to confirm benignity, stability on subsequent CT scans.

Results: 3,318 participants (2,796 male, median age 58.0 years) were included. In 195 participants (5.9%) the initial baseline screen result was adjusted by the radiologist. Adjustment was downwards from positive or indeterminate to negative in two and 119 participants, respectively, and from positive to indeterminate in 65 participants. None of these nodules turned out malignant. In 9/195 participants (4.6%) the screen result was adjusted upwards from negative to indeterminate or indeterminate to positive; two nodules were malignant.

Conclusion: In about one-in-twenty cases of baseline lung cancer screening, nodules were reclassified by the radiologist, leading to reduction of false-positive screen results.

Key points

- The NELSON study allowed radiologists to manually adjust the screen result.
- At baseline, radiologists adjusted the result in about one-in-twenty cases (95.4% downwards).
- Radiologists' adjustments led to reduction of false-positive screen results of 22%.
- Radiologists' expertise can improve nodule classification in addition to a nodule protocol.

Introduction

After publication of the positive results of the National Lung Screening Trial (NLST) [1], interest in lung cancer screening with low-dose chest CT is increasing. All recently published U.S. guidelines recommend lung cancer screening by low-dose chest CT in a high risk population [2–7]. Reducing the rate of false-positive results, however, remains a challenge.

In the Dutch-Belgian randomized lung cancer screening study (Dutch acronym: NELSON study), lung nodule classification was based on nodule volume and nodule growth in terms of volume-doubling time (VDT) [8, 9]. The screen result was classified as negative, indeterminate or positive based on the largest or fastest growing lung nodule. This led to over ten times less false-positive screen results compared to the diameter-based NLST (at baseline 1.7% versus 26.2% false-positives, respectively). In practice, the NELSON study allowed radiologists to manually adjust the screen result at their discretion, for example in case of inappropriate segmentation, or high suspicion of malignancy or benignity without corresponding screen result according to the protocol [8].

To optimize screening efficiency, sensitive lung nodule detection and accurate nodule classification are two important issues [10–12]. Little is known about the impact of readers in lung cancer screening on screening efficiency, in terms of reduction of false-positive results. Therefore, the purpose of this study was to evaluate the impact of radiologists' expertise on test result decisions and accuracy, in particular the impact on lowering false-positive screen results, in a CT lung cancer screening trial.

Methods

Study population

The NELSON multi-centre trial (trial registration number: ISRCTN63545820) was approved by the Dutch Ministry of Health and the ethics board at each participating centre. All participants provided written informed consent. Participants were current and former heavy smokers, aged 50-75 years. Recruitment procedures and selection criteria in the NELSON trial have been published [13]. In total, 15,822 individuals were randomized to no screening (n = 7,907) or screening (n = 7,915) by low-dose chest CT at baseline (1^{st} round), one year later (2^{nd} round), three years later (3^{rd} round) and five-and-a-half years later (4^{th} round), and extra low-dose follow-up CT examinations in case of an indeterminate screening result [14].

For this sub-study, complete data on screen-detected as well as interval lung cancers was obtained via histological specimens of positive screened participants, reassessed by NELSON's chief pathologist, and via linkage with the national cancer registry. Participants randomized to the control group (no screening) were not included in this sub-study (n = 7,907). Participants who did not undergo any CT scan (n = 333), or in whom the baseline screen result was not based on a solid or partial-solid pulmonary nodule (i.e. based on either absence of nodules or presence of a non-solid nodule, n = 4,264) were also excluded from this sub-study. Non-solid nodules were excluded because of the inability

of the semi-automated software used in the NELSON trial to determine volume of these kinds of nodules semi-automatically. Thus, all 3,318 participants in whom the baseline CT screen result was made based on a solid nodule (n = 3,268) or a partial-solid nodule (n = 50) on the baseline CT were included (Figure 9.1). Median age of these participants was 58.0 years (IQR: 55.0-63.0 years), and 2,796 (84.3%) were male. Nodules were followed for up to 6.8 years. Participants had a smoking history with a median of 38.0 pack-years (IQR: 28.0-49.5 years), and 1,876 participants (56.5%) were currently smoking.

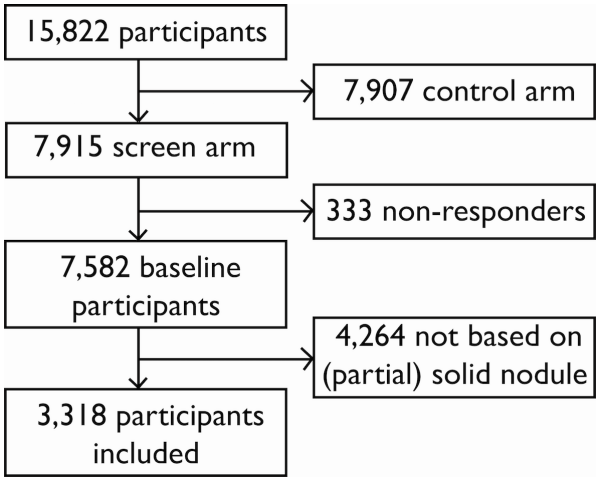


Figure 9.1: Flowchart participants.

CT scanning protocol

At all screening sites 16-row multi-detector CT scanners were used (Sensation-16, Siemens Medical Solutions, Forchheim, Germany, or Mx8000 IDT or Brilliance 16P, Philips Medical Systems, Cleveland, OH, USA). Scans were performed in spiral mode, with 16x0.75mm collimation and 1.5 pitch, in caudo-cranial direction without contrast. Low-dose settings were applied depending on body weight (<50 kg, 50-80 kg and >80 kg), with kVp settings of 80-90 kVp, 120 kVp and 140 kVp, respectively. To achieve CT dose index values of approximately 0.8 mGy, 1.6 mGy and 3.2 mGy, respectively, the mAs settings were adjusted accordingly dependent on the system used. To minimize breathing artefacts, CT scans were performed at maximal inspiration during breath-holding, after appropriate instruction of the subjects. Data were reconstructed at 1.0 mm slice thickness, with 0.7 mm reconstruction increment. Repeat examinations were performed with the same parameters as used at baseline.

Image Reading

All CT images were read twice independently [8, 9]. First readings were performed by one of thirteen radiologists with experience in thoracic CT varying from one year to >20

years. Second readings were done by one of two chest radiologists with at least six years experience. The second readers were full time NELSON readers, reading only low-dose chest CT scans. One of them was trained for three weeks in reading low-dose chest CTs for lung cancer screening at the Department of Radiology Weill Medical College, Cornell University New York, the other was trained by self-education using the ELCAP teaching file. At the moment of reading, the second readers were unaware of the conclusion of the first readers. In case of a discrepancy between the outcome of the first and second reader, a third radiologist with >15 years experience in thoracic CT adjudicated.

For nodule evaluation, the Syngo Lungcare (Leonardo workstation, Somaris/5 VA70C-W, Siemens Medical Solutions) software package for semi-automated pulmonary nodule volume measurements was used in addition to visual evaluation. Baseline and follow-up images were reviewed and displayed simultaneously on one workstation. Lung windows were assessed at a width of 1500 and a level of -500 Hounsfield Units, but the readers were allowed to alter these settings at their discretion.

To measure a nodule, the observer had to mark it with a mouse click. The program then automatically defined the nodule's volume of interest. A three-dimensional template was generated, optimally representing the nodule. A second mouse click initiated the automated volume measurement. In case of unsatisfactory segmentation, manual modification of the segmentation was performed.

Nodule management protocol

The NELSON protocol was described previously [8]. Briefly, screening could lead to three different initial outcomes; a negative screen result (next screening round), an indeterminate result (short-term follow-up examination after three months) or a positive result (referral to a pulmonologist for diagnostic work-up). Work-up, staging, and treatment of participants who were referred to the pulmonologist were according to standard (inter-) national guidelines [8, 15, 16]. At baseline, the screen result was based on nodule size: nodules <50 mm^3 were negative, 50-500 mm^3 indeterminate and nodules >500 mm^3 positive. For partial-solid nodules, the screen result was based on diameter of the subsolid part (<8 mm negative, otherwise indeterminate) and volume of the solid component, similar to solid nodules. At the three-month follow-up CT after baseline, the percentage volume change was calculated for previously detected nodules; <25% volume change was a negative result, $\geq$25% led to the assessment of the VDT. For nodules with VDT <400 days, the final screen result was given as positive, otherwise negative [8].

In practice, the NELSON nodule management protocol, based on semi-automated derived nodule volumes allowed radiologists to overrule screening results that should have been made based according to this protocol. Decisions for manual adjustment of a screen result could for example be made based on high suspicion of malignancy (e.g. enlarged mediastinal lymph nodes), high suspicion of benignity (e.g. nodule with fat component or partially calcified nodules), or inappropriate measurements by the software (e.g. involvement of surrounding structures in case of attached nodules). Nodules demonstrating clearly benign features such as diffuse, central, popcorn or lamellated calcification or internal fat were downwards adjusted, to negative screen result. Other features leading to manual adjustment of screen result were nodule appearance suggestive of scar or fibrosis, in particular plate-like shape, vessel invasion by the nodule, attachment to e.g. fissure or

vessel, and overestimation or underestimation of nodule volume by the LungCare software. Sometimes the nodule features were benign, but due to large volume (>500 mm^3), radiologists were not completely sure about the benign nature of the nodule. Then, the screen result was adjusted from positive to indeterminate, so that growth could be excluded at three-month follow-up CT.

In this study, the final screen result of the baseline screening CT was compared to the screen result of this baseline CT that should have been made according to the NELSON management protocol based on the volume of the largest solid pulmonary nodule. In case of a discrepancy, the CT examination was defined as adjusted, otherwise as non-adjusted. Histology was the reference for diagnosis, or, to confirm benignity, stability of the nodule volume on subsequent CT scans for at least two years after baseline; a period that is considered long enough for regarding if a lung nodule is benign or malignant [17].

Nodule characteristics

Nodules were classified as benign or malignant based on histology, or benign based on stable volume for at least two years after the baseline CT [17, 18]. In addition, nodules were classified based on distance to costal pleura (peripheral or non-peripheral), shape (spherical or non-spherical), and margin (smooth, lobulated, spiculated or irregular) [8, 16, 19]. The distance to costal pleural (only intraparenchymal nodules) was >1/3 of the total hilum-costal pleura distance for non-peripheral nodules and <1/3 for peripheral nodules. A nodule was considered as spherical when its maximum diameter was smaller than twice its minimum diameter; otherwise, it was regarded non-spherical. A nodule was considered as non-smooth when its margin was lobulated, irregular or spiculated, and smooth otherwise [19–21]. Information recorded by the second reader was used. If not available, information recorded by the first reader was used.

Statistical analysis

Parametric data were expressed as mean and 95% confidence interval (95%-CI), non-parametric data as median and interquartile ranges (IQR). Nodule characteristics of adjusted and non-adjusted screen results were compared by a Chi-square test. Partial-solid nodules were analyzed separately from solid nodules because of a different nodule protocol for subsolid nodules [8].

$P \leq 0.05$ was considered to indicate a statistically significant difference. All statistical analyses were performed using SPSS 20.0 (SPSS, Chicago, Ill, USA).

Results

Participants

In 195/3,318 participants (5.9%, 174 male) whose baseline decision was made on a solid or partial-solid lung nodule, the initial baseline CT result that should have been made based on the NELSON nodule management protocol was manually adjusted by the radiologist. In 177 cases (90.8%) the first and second reader agreed on the decision to appoint a

different screen result than that based on semi-automated volumetry. In 18 cases (9.2%) a third reader arbitrated because of difference in screen result between first and second reader. In 17 of 18 cases, the third reader confirmed the decision of the second reader.

Radiologists' adjustments of screen decisions for solid nodules

In Table 9.1, an overview of the protocol adjustments in participants whose baseline screen result was based on a solid nodule is shown. The screen result was manually adjusted downwards from positive to negative or from indeterminate to negative in two and 118 participants, respectively. In 64 participants, the screen result was adjusted downwards from positive to indeterminate. None of these nodules turned out malignant in screening or clinical setting during two years after baseline. In total, the screen result was adjusted downwards in 184 participants (97.4%), resulting in reduction of follow-up CT procedures (n = 118) and reduction of direct referrals to the pulmonologist (n = 66). The screen result was adjusted upwards from negative to indeterminate in one participant. In four participants the screen result was adjusted upwards from indeterminate to positive; two nodules (50%) were diagnosed as lung cancer directly after the baseline CT. This led to a reduction of false-positive rate of 22% (from 131/177 to 66/114).

Table 9.1: Radiologists' adjustments of screen decisions made based on a solid nodule (n = 3,268).

Adjusted result NELSON result	Negative	Indeterminate	Positive
Negative (<50 mm^3)	X	1	-
Indeterminate (50-500 mm^3)	118	X	4 (2x lung cancer baseline)
Positive (>500 mm^3)	2	64	X

Reasons for adjustments in solid nodules

Radiologists had the possibility to report the reason of adjustment of the screen result. They did this in 173 of 189 cases (91.5%). Main reasons for manual adjustment were non-malignant or malignant appearance (n = 95, for example nodule appearance more like a scar or fibrosis, or malignant nodule attachment to vessels), attachment of the nodule to e.g. fissure or vessel (n = 59), and overestimation or underestimation of nodule volume by the LungCare software (n = 14). Please see examples in Figure 9.2.

Adjusted protocol cancer characteristics

The decision to manually adjust the screen result upward in the two baseline lung cancers was based on nodule appearance (vessel invasion by the nodule (stage IIIA adenocarcinoma, Figure 9.3a) and suspicious appearance on baseline CT (no histological diagnosis possible, positive PET-scan (Figure 9.3b), treated with stereotactic radiotherapy). Baseline volumes of these nodules were 245 mm^3 and 361 mm^3 (Figure 9.3).

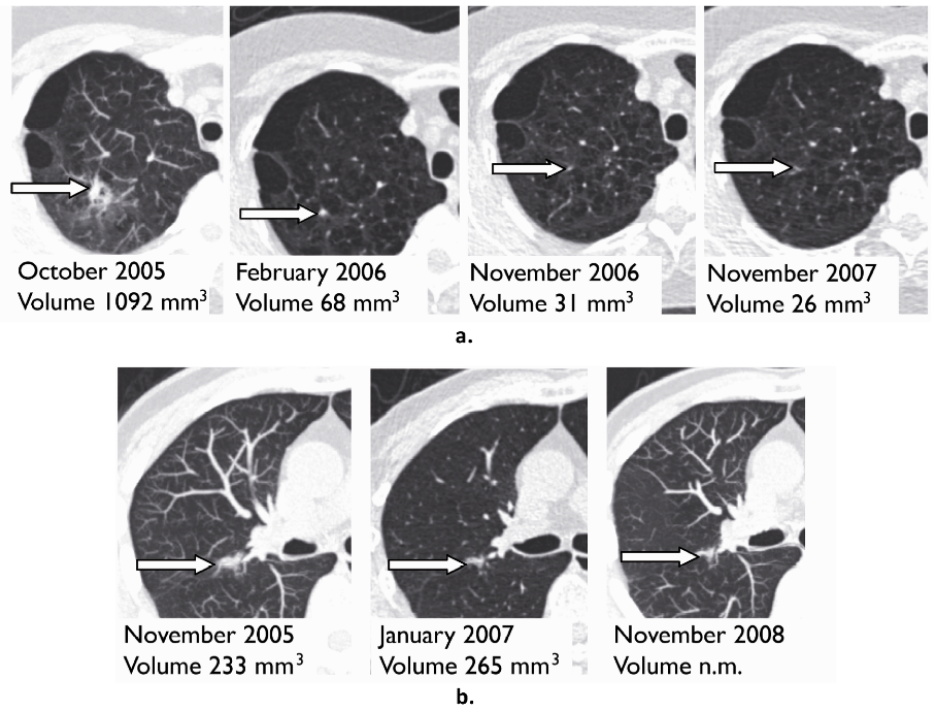


Figure 9.2: Examples of downward classified benign lesions. Downwards adjusted lesion in a 54-year-old female (a) and a 72-year-old male (b). Axial computed tomography (CT) shows lesions with baseline volumes of 1092 mm^3 (a) and 233 mm^3 (b). Reasons for adjustment were appearance as fibrosis or scar making the radiologist to adjust the screen result from positive to indeterminate to see what happens in three months (a), and attachment to the fissure and appearance more like fibrosis, making the radiologist to adjust the screen result from indeterminate to negative (b).

Nodule characteristics

The characteristics of the largest solid nodule present in the participants, with downwards and without screen result adjustment, are compared in Table 9.2. In participants whose baseline CT screen result was manually adjusted downwards, nodules were more often located non-peripheral than peripheral (P<0.01), compared to participants without screen result adjustment. In downwards adjusted screen results, nodule margin was relatively less often smooth, and more often irregular, mostly with high suspicion of fibrosis based on appearance, compared to nodules in non-adjusted screen results (P<0.001). No statistically significant difference in nodule shape was found between nodules in adjusted and non-adjusted screen results.

In five participants with upwards screen result adjustment, characteristics of the largest solid nodule were as follows: three nodules were located peripherally, one non-peripherally, and one non-intraparenchymally; all five nodules were non-spherical; one nodule had a lobulated margin, two were spiculated, and two were irregular.

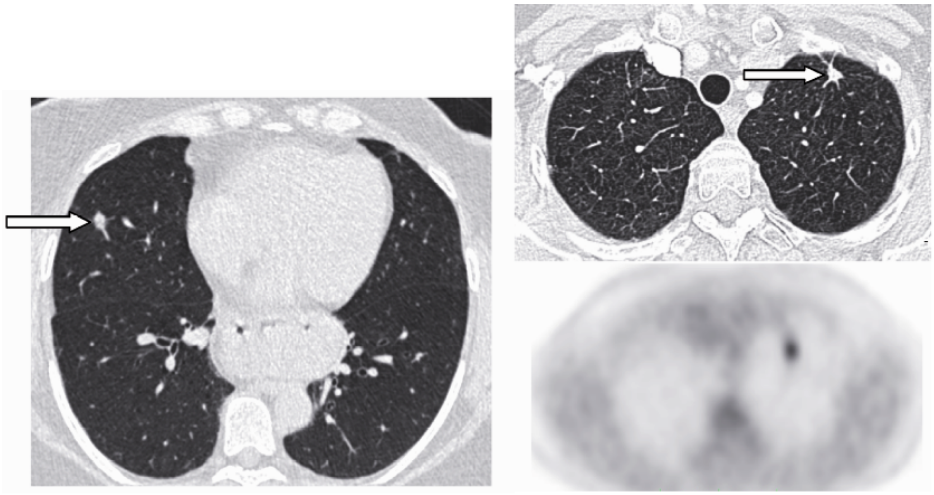


Figure 9.3: Examples of upwards classified malignant lesions. Upwards adjusted lesion in a 60-year-old female (a) and a 57-year-old female (b). Axial CT shows an irregular nodule in the right middle lobe with volume of 245 mm^3(a, adenocarcinoma), and a spiculated nodule in the left upper lobe with volume of 361 mm^3(b, no histological diagnosis possible). The FDG-PET scan shows uptake in the lung nodule in (b).

Table 9.2: Characteristics of solid nodules of downward adjusted and non-adjusted screen results.

	Total	Downward Adjusted	Non-adjusted	*P*-value
Total	3263	184 (5.6)	3079 (94.4)	
Distance to costal pleural[a,b]				
Peripheral	2018	100 (5.0)	1918 (95.0)	0.008
Non-peripheral	611	47 (7.7)	564 (92.3)	
Shape[c]				
Spherical	733	45 (6.1)	688 (93.9)	0.346
Non-spherical	2450	139 (5.7)	2311 (94.3)	
Margin[b,d]				
Smooth	2413	97 (4.0)	2316 (96.0)	<0.001
Lobulated	555	38 (6.8)	517 (94.2)	
Spiculated	154	17 (11.0)	137 (89.0)	
Irregular	87	27 (31.0)	60 (69.0)	

Unless otherwise indicated, data are numbers of nodules, with percentages in parentheses.
[a] 2633 nodules were located intraparenchymal.
[b] Indicates a statistically significant difference (Chi-square).
[c] 80 cases (2.4%) missing.
[d] 54 cases (1.7%) missing.

Partial-solid nodules

In 50 participants, the baseline screen decision was based on a partial-solid lung nodule. In six participants (12%), the screen result was adjusted manually. Adjustment was made

downwards, from indeterminate to negative, in one participant (volume overestimation by the software), and from positive to indeterminate in one participant (nodule appearance more like fibrosis). In four participants, the result was adjusted upwards, from negative to indeterminate. Reason for these upwards adjustment was the large nodule diameter (Table 9.3).

Table 9.3: Radiologists' adjustments of screen decisions made based on a partial-solid nodule (n = 50).

Adjusted result NELSON result	Negative	Indeterminate	Positive
Negative (solid part <50 mm^3)	X	4	-
Indeterminate (solid part 50-500 mm^3)	1	X	-
Positive (solid part >500 mm^3)	-	1	X

Discussion

To the best of our knowledge, this study is the first report on manual protocol adjustments by radiologists in lung cancer screening by low-dose CT. We found that in participants whose baseline screen result was based on a solid lung nodule, 5.9% of the baseline CT screen results based on the NELSON nodule management protocol were adjusted by the radiologist. About 95% of screen results were adjusted downwards. In total, this led to 5.6% less short-term follow-up CT procedures or referrals after the baseline CT, while none of the nodules turned out to be cancer during two years after baseline.

Noncalcified lung nodules are found in up to 69% of participants in lung cancer screening. However, even in this group of persons with a high risk of developing lung cancer, only 1.0-3.6% of these nodules are diagnosed as lung cancer [1, 22]. In lung cancer screening, it is of major importance to minimize the rate of false-positive test-results, and thus increase specificity, without missing lung cancer cases, especially when lung cancer screening by low-dose CT becomes more widespread. Optimization of the specificity is needed to reduce negative psychological effects [23, 24], costs and harm from unnecessary invasive procedures and radiation exposure. This study showed that that radiologists' expertise could be beneficial to reduce the false-positive rate. Considering the large population eligible for lung cancer screening, a rate of 5.6% downward adjustments may lead to a considerable decrease in follow-up CTs and (invasive) workup.

The radiologists' decisions for manual adjustment were most often based on nodule appearance or nodule attachment. Research published after the baseline screening round of the NELSON study showed that vessel or fissural attached nodules have very low probability of malignancy [16, 25]. De Hoop et al. showed that perifissural nodules are rarely malignant, even in case of rapid growth [26]. In this study, we also found that in two years after baseline, none of the downward adjusted nodules turned out to be lung cancer.

Adjustments of the screen result were performed more often in case of non-peripherally

located nodules and in case of spiculated and irregular nodules. Previous studies have shown that certain benign lesions (i.e. fibrosis), as well as malignant nodules, have spiculated or irregular margins [27, 28]. Therefore, the interference of the radiologist on screen result is more pronounced in lesions with complex morphology. In particular, a plate-like appearance of an irregular or spiculated nodule led the readers to interpret the nodule as fibrosis instead of suspicious for lung cancer.

The radiologists who read the CT examinations in the NELSON study had experience in reading thoracic CT varying from one year to >20 years. The first and second readers had experience ranging from one to 10 years. Even so, in over 90% of cases, the first and second reader agreed in adjusting the screen result. Only in 9.2% of the adjusted screen results, a third, expert radiologist with >20 years of experience arbitrated. Therefore, we do not think that years of experience in reading chest CT significantly influenced the decisions made for adjustments.

The median follow-up time of the 195 nodules on which the screen result adjustment was based was 5.5 years. In the third screening round, three years after baseline, lung cancer was detected in two participants whose baseline CT result was manually adjusted from positive to indeterminate. Both nodules on which the manual adjustment at baseline was based were at time of diagnosis stage I adenocarcinoma. In both cases, radiologists adjusted the screen result based on appearance (one because the nodule was not a typical nodule, the other because the nodule looked more like a scar) to an indeterminate result, leading to a short-term follow-up examination. Already after the short-term follow-up CT in the baseline round, both participants were referred to a pulmonologist because of growth. At that time, work-up by the pulmonologist according to (inter)national guidelines turned out negative in both cases. Thus, these nodules were concluded not to concern lung cancer. Later, these nodules evolved into stage I lung cancers, and were detected by screening three years after baseline. The aim of the stringent NELSON nodule management protocol is watchful waiting with follow-up CT procedures for indeterminate nodules and solely referral for large or fast growing lung nodules. As both lung cancers diagnosed in year three had a negative work-up in the first year, and were in stage I at moment of diagnosis, by definition, we do not consider them to be missed lung cancers.

One limitation of this study is the inability of the Syngo LungCARE software package to calculate the volume of sub-solid nodules (1.9% of all non-calcified nodules). In the main part of this study, we therefore included only solid nodules. For the vast majority of solid nodules, semi-automated measurements were found to be highly reproducible [29]. In 86% of 4,225 screen-detected solid nodules evaluated in the NELSON substudy by Wang et al, double reading obtained complete agreement in volume. Volume differences >15% were found in only 4% of nodules [29]. When measured volume differed between first and second reader, results from the second reader were used for further analyses. The software was able to calculate the solid part of partial-solid nodules (0.7% of all solid nodules), and we provided an overview of nodule adjustments in participants whose screen result was based on a partial-solid nodule ($n = 50$). Since the LungCARE software was not able to semi-automatic measure the subsolid part of the partial-solid nodule, nodule management based on nodule volume of the solid part alone may be insufficient. It is expected that separate guidelines will be developed for the management of sub-solid nodules detected in lung cancer screening, such as recently published for incidentally detected sub-solid

nodules [30].

In conclusion, in baseline lung cancer screening, readers adjusted screen results in about one-in-twenty cases (95.4% downwards), resulting in lowering of follow-up CT procedures (n = 119) or direct referrals to the pulmonologist (n = 67), and a false-positive reduction of 22%. Therefore, radiologists' expertise can improve nodule classification in addition to a volume-based nodule management protocol.

References

[1] National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England journal of medicine*, 365(5):395–409.

[2] US Preventive Services Task Force (2013). Screening for Lung Cancer: U.S. Preventive Services Task Force Recommendation Statement. URL `http://www.uspreventiveservicestaskforce.org/draftrec.htm`.

[3] Jaklitsch MT, Jacobson FL, Austin JH, Field JK, Jett JR, Keshavjee S, et al. (2012). The American Association for Thoracic Surgery guidelines for lung cancer screening using low-dose computed tomography scans for lung cancer survivors and other high-risk groups. *The Journal of thoracic and cardiovascular surgery*, 144(1):33–38.

[4] Detterbeck FC, Mazzone PJ, Naidich DP, and Bach PB (2013). Screening for lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, 143(5 Suppl):e78S–92S.

[5] Wender R, Fontham ET, Barrera Jr E, Colditz GA, Church TR, Ettinger DS, et al. (2013). American Cancer Society lung cancer screening guidelines. *CA: a cancer journal for clinicians*, 63(2):107–117.

[6] American Lung Association (2012). Providing Guidance on CT Lung Cancer Screening to Patiens and Physicians. URL `http://www.lung.org/lung-disease/lung-cancer/lung-cancer-screening-guidelines/lung-cancer-screening.pdf`.

[7] Humphrey LL, Deffebach M, Pappas M, Baumann C, Artis K, Mitchell JP, et al. (2013). Screening for Lung Cancer With Low-Dose Computed Tomography: A Systematic Review to Update the U.S. Preventive Services Task Force Recommendation. *Annals of Internal Medicine*, 159(6):411–20.

[8] Xu DM, Gietema H, de Koning H, Vernhout R, Nackaerts K, Prokop M, et al. (2006). Nodule management protocol of the NELSON randomised lung cancer screening trial. *Lung cancer*, 54(2):177–184.

[9] van Klaveren RJ, Oudkerk M, Prokop M, Scholten ET, Nackaerts K, Vernhout R, et al. (2009). Management of lung nodules detected by volume CT scanning. *The New England journal of medicine*, 361(23):2221–2229.

[10] Christe A, Leidolt L, Huber A, Steiger P, Szucs-Farkas Z, Roos JE, et al. (2013). Lung cancer screening with CT: Evaluation of radiologists and different computer assisted detection software (CAD) as first and second readers for lung nodule detection at different dose levels. *European Journal of Radiology*, 82(12):e873–8.

[11] Zhao Y, de Bock GH, Vliegenthart R, van Klaveren RJ, Wang Y, Bogoni L, et al. (2012). Performance of computer-aided detection of pulmonary nodules in low-dose CT: comparison with double reading by nodule volume. *European radiology*, 22(10):2076–2084.

[12] Kakinuma R, Ashizawa K, Kobayashi T, Fukushima A, Hayashi H, Kondo T, et al. (2012). Comparison of sensitivity of lung nodule detection between radiologists and technologists on low-dose CT lung cancer screening images. *The British journal of radiology*, 85(1017):e603–8.

[13] van Iersel CA, de Koning HJ, Draisma G, Mali WP, Scholten ET, Nackaerts K, et al. (2007). Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *International journal of cancer.Journal international du cancer*, 120(4):868–874.

[14] Horeweg N, van der Aalst CM, Vliegenthart R, Zhao Y, Xie X, Th Scholten E, et al. (2013). Volumetric computer tomography screening for lung cancer: three rounds of the NELSON trial. *The European respiratory journal*, 42(6):1659–1667.

[15] CBO (2004). *Guideline - non-small cell lung carcer: staging and treatment*. Van Zuiden Communications BV, Alphen aan de Rijn, The Netherlands.

[16] Xu DM, van der Zaag-Loonen HJ, Oudkerk M, Wang Y, Vliegenthart R, Scholten ET, et al. (2009). Smooth or attached solid indeterminate nodules detected at baseline CT screening in the NELSON study: cancer risk during 1 year of follow-up. *Radiology*, 250(1):264–272.

[17] MacMahon H, Austin JH, Gamsu G, Herold CJ, Jett JR, Naidich DP, et al. (2005). Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology*, 237(2):395–400.

[18] Gould MK, Donington J, Lynch WR, Mazzone PJ, Midthun DE, Naidich DP, et al. (2013). Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, 143(5 Suppl):e93S–120S.

[19] Gurney JW, Lyddon DM, and McKay JA (1993). Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis. Part II. Application. *Radiology*, 186(2):415–422.

[20] Takashima S, Sone S, Li F, Maruyama Y, Hasegawa M, Matsushita T, et al. (2003). Small solitary pulmonary nodules ($<$ or $=$1 cm) detected at population-based CT screening for lung cancer: Reliable high-resolution CT features of benign lesions. *American journal of roentgenology*, 180(4):955–964.

[21] Zhao YR, Heuvelmans MA, Dorrius MD, van Ooijen PM, Wang Y, de Bock GH, et al. (2014). Features of Resolving and Nonresolving Indeterminate Pulmonary Nodules at Follow-up CT: The NELSON Study. *Radiology*, 270(3):872–879.

[22] Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS, et al. (2002). CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. *American journal of roentgenology*, 178(5):1053–1057.

[23] Brewer NT, Salz T, and Lillie SE (2007). Systematic review: the long-term effects of false-positive mammograms. *Annals of Internal Medicine*, 146(7):502–510.

[24] van den Bergh KA, Essink-Bot ML, Bunge EM, Scholten ET, Prokop M, van Iersel CA, et al. (2008). Impact of computed tomography screening for lung cancer on participants in a randomized controlled trial (NELSON trial). *Cancer*, 113(2):396–404.

[25] Ahn MI, Gleeson TG, Chan IH, McWilliams AM, Macdonald SL, Lam S, et al. (2010). Perifissural nodules seen at CT screening for lung cancer. *Radiology*, 254(3):949–956.

[26] de Hoop B, van Ginneken B, Gietema H, and Prokop M (2012). Pulmonary Perifissural Nodules on CT Scans: Rapid Growth Is Not a Predictor of Malignancy. *Radiology*, 265(2):611–616.

[27] Li F, Sone S, Abe H, Macmahon H, and Doi K (2004). Malignant versus benign nodules at CT screening for lung cancer: comparison of thin-section CT findings. *Radiology*, 233(3):793–798.

[28] Zwirewich CV, Vedal S, Miller RR, and Muller NL (1991). Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation. *Radiology*, 179(2):469–476.

[29] Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. (2008). Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. *Radiology*, 248(2):625–631.

[30] Naidich DP, Bankier AA, MacMahon H, Schaefer-Prokop CM, Pistolesi M, Goo JM, et al. (2013). Recommendations for the management of subsolid pulmonary nodules detected at CT: a statement from the Fleischner Society. *Radiology*, 266(1):304–317.